

AMES

NASA Technical Memorandum 88371

Smoothing and the Second Law

Marshal L. Merriam

(NASA-TM-88371) SMOOTHING AND THE SECOND
LAW (NASA) 23 p Avail: NTIS HC A02/MF
A01 CSCI 12A

N87-24146

H1/64 Unclas
0080155

September 1986

NASA
National Aeronautics and
Space Administration

Smoothing and the Second Law

Marshal L. Merriam, Ames Research Center, Moffett Field, California

September 1986



National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035

Smoothing and the Second Law

Marshal L. Merriam¹

NASA Ames Research Center
Moffett Field, California

Abstract

The technique of obtaining second order, oscillation free, total variation diminishing (TVD), scalar difference schemes by adding a limited diffusive flux ("smoothing") to a second order centered scheme is explored. It is shown that such schemes do not always converge to the correct physical answer. The approach presented here is to construct schemes that numerically satisfy the second law of thermodynamics on a cell by cell basis. Such schemes can only converge to the correct physical solution and in some cases can be shown to be TVD. An explicit scheme with this property and second order spatial accuracy was found to have an extremely restrictive time step limitation ($\Delta t < \Delta x^2$). Switching to an implicit scheme removed the time step limitation.

1. Introduction

The numerical simulation of fluid flow is an interesting problem. For many cases of interest the equations and constitutive relations are well known and have been for some time. A great deal of excellent work has been devoted to solving these equations. The current state of the art for convection dominated problems is represented by various total variation diminishing (TVD) schemes and the closely related essentially non-oscillatory (ENO) schemes. These schemes can be constructed using geometric arguments. Unfortunately, it is possible to construct wiggle free schemes that give physically impossible answers. Another approach, presented here, is to construct schemes that numerically satisfy the second law of thermodynamics on a cell by cell basis. Such schemes do not allow unphysical solutions and in some special cases can be shown to be TVD.

Any general numerical method for the simulation of fluid flow must deal with the case of convection dominated flows with shock waves. Such cases are highly nonlinear and represent a real challenge to many algorithms. This work deals with such problems.

Recently Sweby [1] proposed a method for constructing TVD schemes. It was the investigation of this method that led to the work of this paper, which shows how to construct entropy satisfying, second order accurate, central difference schemes stabilized by a limited diffusive flux. In some cases these can be shown to be TVD.

All the schemes reviewed by Sweby (Roe[2], Van Leer[3], Osher-Chakravarthy[4]) were based on the second order Lax-Wendroff scheme [5] and consequently had the usual drawbacks of such a scheme. Specifically, such schemes have a steady state solution that depends on the time step and are, therefore, unsuitable for implicit methods. Also, not all TVD

Presented at First World Congress on Computational Mechanics, Austin, Texas, September 22-25, 1986.

¹Research Scientist, Computational Fluid Dynamics Branch.

schemes converge to the correct solution. It is possible to construct, for the wave equation, a TVD scheme that violates entropy, thus resulting in a square wave for arbitrary initial data. The objective of this work is to construct a second order accurate space difference scheme that satisfies a cell entropy inequality and converges to a solution that is independent of the time step.

For entropy satisfying schemes using explicit Euler time advance, extremely restrictive time step limitations are needed to achieve second order accuracy. On the other hand, using implicit Euler time advance, a scheme can be constructed that has all the desired properties and no time step limitation.

The approach presented here is still unproven but shows great promise. It is one dimensional though the extension to multidimensions is straightforward. Although it has been shown (Goodman and LeVeque [6]) that TVD schemes are at most first order accurate in multiple space dimensions, it appears that entropy satisfying schemes do not share this restriction. Furthermore they are, in principle, extensible to the Navier-Stokes equations with finite rate chemistry.

2. Generalized Entropy and Convection Processes

Convection processes are often described in terms of control volumes. For a fixed control volume, a general convection process can be written

$$(2.1) \quad \int_{\mathcal{V}} q_{,t} d\mathcal{V} + \oint_S \mathbf{f}(\mathbf{q}) \cdot d\mathbf{n} = 0$$

where the integrals are over the volume and over the surrounding surface respectively. The vector $d\mathbf{n}$ is the usual outward facing unit normal. For the case of gasdynamics, this covers conservation of mass, momentum and energy. Satisfaction of the first law of thermodynamics follows as a direct consequence. Unfortunately Eq. 2.1 says nothing about the second law of thermodynamics and admits unphysical solutions such as expansion shocks.

For uniqueness, the second law of thermodynamics must also be satisfied. This can also be stated for a control volume.

$$(2.2) \quad \int_{\mathcal{V}} S(q)_{,t} d\mathcal{V} + \oint_S \mathbf{F}(q) \cdot \mathbf{n} \geq 0$$

In this equation $S(q)$ is the generalized entropy, a scalar function of q , and $\mathbf{F}(q)$ is the so called generalized entropy flux. The quantity $\mathbf{F}(q)$ is properly a vector in the sense that it has a magnitude and a direction but not in the sense of q which contains several independent quantities. In one dimension, no generality is lost by representing F as a scalar quantity. Though not limited to the thermodynamic definitions, S and F must be chosen to satisfy the following relations.

$$(2.3) \quad S_{,qq} < 0$$

$$(2.4) \quad F_{,q} = S_{,q} f_{,q}$$

The comma notation used here denotes differentiation. In the case where the independent variable q is a vector of unknowns, Eq. 2.3 is taken to mean that $S_{,qq}$ is a negative definite matrix. If q is a vector with more than two unknowns, the existence of S and F is not guaranteed. In any case, S may not be unique. For the case of the Euler equations it can be verified by substitution that the thermodynamic entropy satisfies these equations. Further discussion of generalized entropy can be found in a classic paper by Lax [7]. The numerical analog of Eqs. 2.2-2.4 will be referred to as the entropy condition.

There is one other requirement. The fluxes $f(q)$ must remain bounded everywhere. For example, the Kutta condition is a way of satisfying this requirement near an airfoil trailing edge. The numerical analog of this requirement will be called the flux consistency condition. The flux consistency condition used in this work is

$$(2.5) \quad (f(q_{j+1}) - f(q_{j+\frac{1}{2}}))(f(q_{j+\frac{1}{2}}) - f(q_j)) \geq 0$$

which requires that extrema in the flux occur at grid points. This is slightly stronger than the consistency condition suggested by Lax [7]. Lax suggests

$$(2.6) \quad f(q_{j+1}) = f(q_{j+\frac{1}{2}}) = f(q_j) \quad \text{if} \quad q_j = q_{j+1}$$

which amounts to consistency in a Taylor series sense. The value of a flux consistency condition is twofold. The first is accuracy. Satisfaction of Eq. 2.6 in a stable scheme is enough to ensure convergence to the correct weak solution as the grid is refined in space and time. The second is entropy satisfaction. Satisfaction of Eq. 2.5 is needed to ensure that the entropy inequalities are satisfied between gridpoints.

For various historical reasons, and because the notation is more compact, Eqs. 2.1 and 2.2 are often replaced by an equivalent partial differential equation (PDE). For example Eq. 2.1 is equivalent to the PDE

$$(2.7) \quad q_{,t} + f(q)_{,x} = 0$$

in one dimension. The remainder of this work will use the simpler PDE notation, though in principle the control volume notation could have been used throughout.

This work deals with numerical satisfaction of the entropy condition and the flux consistency condition. These often ignored areas are important because they make the solution unique, and because they are the qualitative reason given for artificial dissipation or "smoothing."

3. Fun with the Wave Equation

The simplest convection model available is the scalar wave equation.

$$(3.1) \quad u_{,t} + cu_{,x} = 0 \quad c > 0$$

This linear equation has a simple analytic solution and does not support shock waves in the usual sense. In spite of this simplicity, it is not trivial to produce a good numerical algorithm for the wave equation. A good numerical algorithm should be conservative, second order accurate in space, and should satisfy a cell entropy for any CFL number sufficiently small.

The simplest second order scheme is explicit Euler in time, central differenced in space.

$$(3.2) \quad u^j = u_j - \frac{\nu}{2}(u_{j+1} - u_{j-1})$$

$$\nu \equiv \frac{c\Delta t}{\Delta x}$$

In this notation (which is only suitable for schemes with two time levels) superscripts indicate values at the new time level and subscripts represent values at the old time level. In the notation of Sweby this may be written

$$(3.3) \quad u^j = u_j - \nu\Delta u_{j-\frac{1}{2}} - \Delta^-(\frac{1}{2}\nu\Delta u_{j+\frac{1}{2}})$$

where we use the convention

$$\Delta^+ u_j = \Delta u_{j+\frac{1}{2}} = \Delta^- u_{j+1} = u_{j+1} - u_j$$

As written, the scheme is known to be unconditionally unstable. Traditionally one adds a second order dissipation term.

$$(3.4) \quad u^j = u_j - \nu\Delta u_{j-\frac{1}{2}} - \Delta^-(\frac{1}{2}\nu\Delta u_{j+\frac{1}{2}}) + \Delta^-(\frac{1}{2}\nu\alpha_{j+\frac{1}{2}}\Delta u_{j+\frac{1}{2}})$$

The variable $\alpha_{j+\frac{1}{2}}$ is a smoothing coefficient which may or may not vary from point to point. Combining terms and using $\phi_{j+\frac{1}{2}} \equiv 1 - \alpha_{j+\frac{1}{2}}$ gives

$$(3.5) \quad u^j = u_j - \nu\Delta u_{j-\frac{1}{2}} - \Delta^-(\frac{1}{2}\nu\phi_{j+\frac{1}{2}}\Delta u_{j+\frac{1}{2}})$$

If $\phi_{j+\frac{1}{2}}$ is a constant, the scheme that results is the standard central difference scheme with second order artificial viscosity as proposed by Lax [7]. The coefficient of artificial

viscosity is α in this case. If $\phi_{j+\frac{1}{2}} = 0$, the first order upwind scheme results. If $\phi_{j+\frac{1}{2}} = 1$, Eq. 3.2 is recovered. If $\phi_{j+\frac{1}{2}} = 2$, a first order downwind scheme results, suitable if $c < 0$.

All constant values of $\phi_{j+\frac{1}{2}}$, except $\phi_{j+\frac{1}{2}} = 0$, result in schemes with spurious oscillations, otherwise known as wiggles. A wiggle free, second order scheme will require nonconstant values of $\phi_{j+\frac{1}{2}}$. One way to produce such a scheme is with the TVD ideas formalized by Harten and explained by Sweby.

Expanding the last term of Eq. (3.5) gives

$$(3.6) \quad u^j = u_j - \nu \Delta u_{j-\frac{1}{2}} - \frac{1}{2} \nu \phi_{j+\frac{1}{2}} \Delta u_{j+\frac{1}{2}} + \frac{1}{2} \nu \phi_{j-\frac{1}{2}} \Delta u_{j-\frac{1}{2}}$$

The standard practice at this point is to introduce the quantity $r_j \equiv \frac{\Delta u_{j-\frac{1}{2}}}{\Delta u_{j+\frac{1}{2}}}$ and write Eq. 3.6 as

$$(3.7) \quad u^j = u_j - \nu \Delta u_{j-\frac{1}{2}} - \frac{1}{2} \nu \frac{\phi_{j+\frac{1}{2}}}{r_j} \Delta u_{j-\frac{1}{2}} + \frac{1}{2} \nu \phi_{j-\frac{1}{2}} \Delta u_{j-\frac{1}{2}}$$

This allows collection of terms so that Harten's TVD conditions [8] can be applied.

$$(3.8) \quad u^j = u_j - \nu \left[1 + \frac{1}{2} \left(\frac{\phi_{j+\frac{1}{2}}}{r_j} - \phi_{j-\frac{1}{2}} \right) \right] \Delta u_{j-\frac{1}{2}}$$

The TVD proof says that the scheme will be TVD if we can insure that

$$(3.9) \quad 0 \leq \nu \left[1 + \frac{1}{2} \Phi_j \right] \leq 1$$

where $\Phi_j \equiv \left(\frac{\phi_{j+\frac{1}{2}}}{r_j} - \phi_{j-\frac{1}{2}} \right)$. With a little algebra the expression can be written (for $\nu > 0$) as

$$(3.10) \quad -2 \leq \Phi_j \leq \frac{2(1-\nu)}{\nu}$$

Inequalities (3.10) are plotted in Fig. 1. Everything in the shaded region is TVD. Generally, Φ_j takes on both positive and negative values in the neighborhood of zero. The object is to pick a maximum CFL number such that a reasonable range of Φ_j can be used. Since Φ_j is independent of ν the object is to pick rectangular regions from the shaded region.

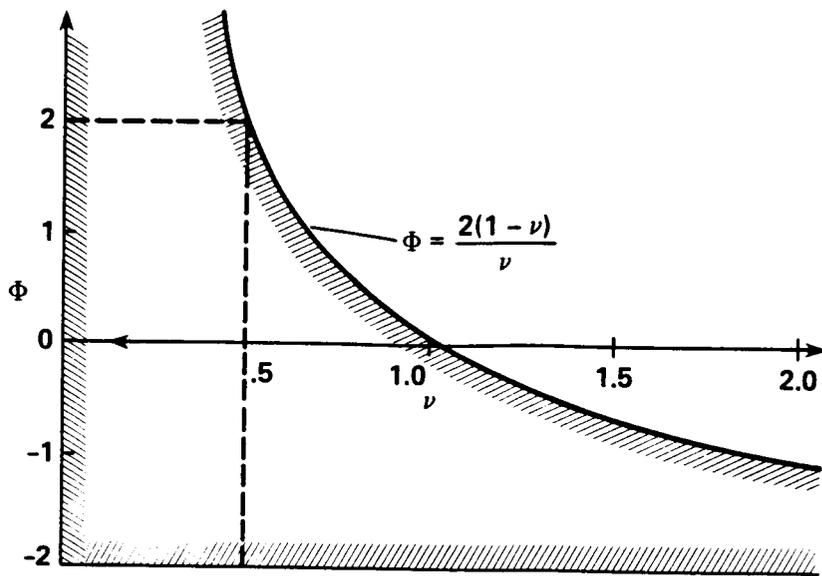


Figure 1. Schemes in the shaded region are TVD.

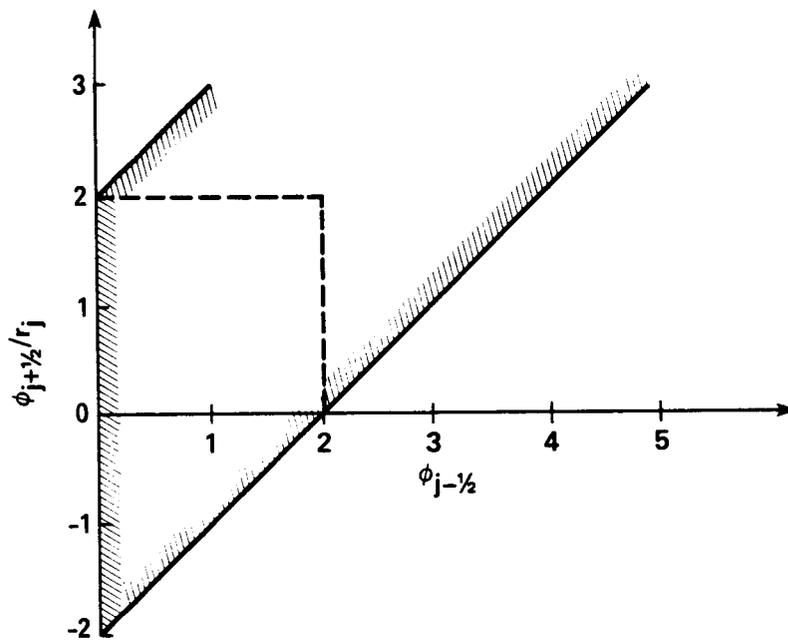


Figure 2. Schemes in the shaded region are TVD for $\nu \leq \frac{1}{2}$.

For example, if $\nu \gg 1$ then $\Phi_j = -2$. Unfortunately, this leads to an inconsistent scheme, $u_{j+1} = u_j$, which is stable and TVD but gives the wrong answer. Consistent

schemes (in the sense of Eq. 2.5) are possible if $\nu \leq 1$. If $\nu = \frac{1}{2}$ the constraints are

$$(3.11) \quad -2 \leq \frac{\phi_{j+\frac{1}{2}}}{r_j} - \phi_{j-\frac{1}{2}} \leq 2$$

where we have substituted in the definition for Φ_j . Generally, the coefficients $\phi_{j+\frac{1}{2}}$ and $\phi_{j-\frac{1}{2}}$ should be picked independently for algorithmic simplicity. Assuming that $\phi_{j+\frac{1}{2}}$ can be chosen as a function of r_j , inequalities 3.11 can be satisfied within the shaded region of Fig. 2. Negative values of $\phi_{j-\frac{1}{2}}$ imply schemes with more dissipation than the first order upwind scheme and are excluded.

Once again the task is to pick a suitable rectangular region from the shaded part of Fig. 2. Except for extrema it is easily shown that $r_j \approx 1$, with equality being achieved in the limit as $\Delta x \rightarrow 0$. Since all numerical analysis is based on this limiting case, we had better include it in our plans. A similar argument holds that except for discontinuities, $r_{j-1} \approx r_j$ so that $\phi_{j+\frac{1}{2}} \approx \phi_{j-\frac{1}{2}}$. This being the case most of the action will occur on a line of slope one through the origin. The rectangular region with the largest amount of this line is the square defined by

$$(3.12) \quad 0 \leq \left(\frac{\phi_{j+\frac{1}{2}}}{r_j}, \phi_{j-\frac{1}{2}} \right) \leq 2$$

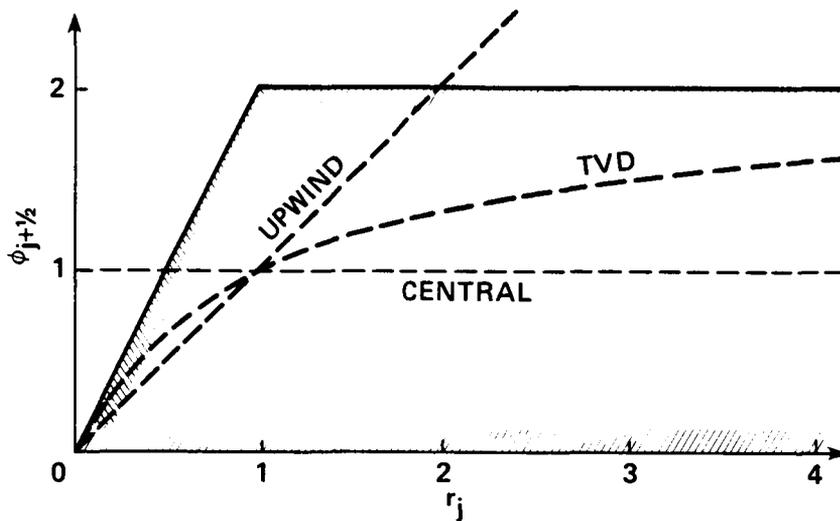


Figure 3. Schemes in which $\phi(r)$ is within the shaded region are TVD.

It is important to see that these inequalities must be satisfied at each point and that $\phi_{j+\frac{1}{2}}$ will replace $\phi_{j-\frac{1}{2}}$ at the point $j+1$. Thus, $\phi_{j-\frac{1}{2}}$ can be replaced with $\phi_{j+\frac{1}{2}}$ in Eq. (3.12) and the resulting inequalities plotted in Fig. 3. The entire line $\phi_{j+\frac{1}{2}} = 0$ is included in the shaded region. Notice that $\phi = 0$ represents the only choice for a constant coefficient TVD scheme of this type, and it is first order accurate. The idea here is to construct some function $\phi(r)$ that lies in the shaded region and results in a second order scheme. The function $\phi = 1$ yields the standard central difference scheme with no dissipation. The function $\phi = r$ yields a three point, second order upwind scheme. Neither of these is stable without dissipation and neither is TVD. Notice that both schemes pass through the point $\phi = r = 1$. Recall that $r \rightarrow 1$ as $\Delta x \rightarrow 0$ for all schemes. For accuracy it is desirable that $\phi(1) = 1$ so that the dissipation will vanish as the mesh is refined. All functions $\phi(r)$ that pass through this point will be second order except at extrema and discontinuities. One function that lies between the two linear second order schemes and passes through the point $\phi = r = 1$ is shown in Fig. 3. Its formula is

$$(3.13) \quad \begin{aligned} \phi_{j+\frac{1}{2}} &= 0, & r_j < 0 \\ \phi_{j+\frac{1}{2}} &= \frac{2r_j}{r_j + 1}, & r_j \geq 0 \end{aligned}$$

This limiter is attributed to van Leer, though intended for use on Lax-Wendroff type TVD schemes. This scheme is programmable and vectorizable, forming successively r_j from the definition, $\phi_{j+\frac{1}{2}}$ from Eq. 3.13, and finally u^j from Eq. 3.5.

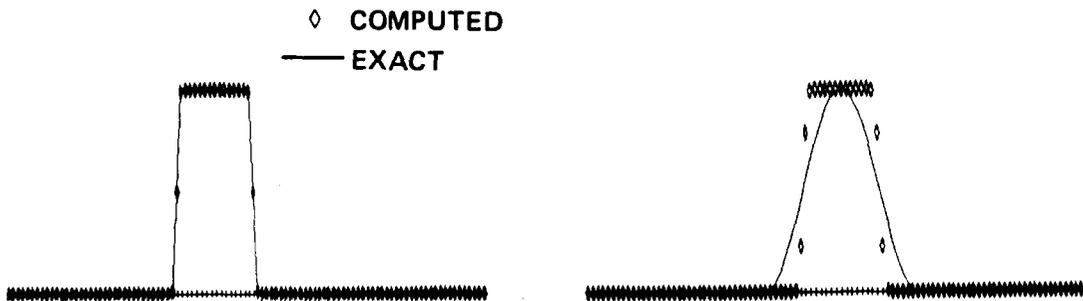


Figure 4. Results for a conservative TVD scheme that violates entropy.

Figure 4 shows the performance of this simple TVD scheme. On the square wave initial data at a CFL number of $\frac{1}{2}$ the solution is nearly perfect, as good as the mesh will permit. The reason is clear when the case of sine wave initial data is examined. Here too, the solution is a square wave. In fact, for an arbitrary initial pulse this scheme will eventually produce a convecting square wave.

Intuitively a scheme that produces nothing but square waves seems to be a bad scheme. What makes it a bad scheme? Not conservation, for u is conserved exactly (the area under

the pulse is unchanged. Not wiggles, the scheme is TVD, as advertised. It is the failure of this scheme to satisfy the entropy inequality, important even in the absence of shock waves and discontinuities, that results in unphysical solutions.

4. Entropy and the Wave Equation

The definitions for entropy and the associated entropy flux are not unique. One such pair is $S = -u^2$ and $F = -cu^2$. It can be verified by substitution that this satisfies Eqs. 2.3 and 2.4 as required.

The actual entropy inequality Eq. (2.2) can only be evaluated if something is known about how $S(u)$ varies over the control volume. In this work u is assumed to be piecewise constant so $S(u)$ is constant over the control volume. Under this assumption $u_{j+\frac{1}{2}}$ is undefined, taking on any value between u_j and u_{j+1} .

Using the explicit Euler time advance as before gives the difference scheme

$$(4.1) \quad u^j = u_j - \nu(u_{j+\frac{1}{2}} - u_{j-\frac{1}{2}})$$

and the corresponding discrete entropy statement

$$(4.2) \quad (u^j)^2 - u_j^2 + \nu(u_{j+\frac{1}{2}}^2 - u_{j-\frac{1}{2}}^2) \leq 0$$

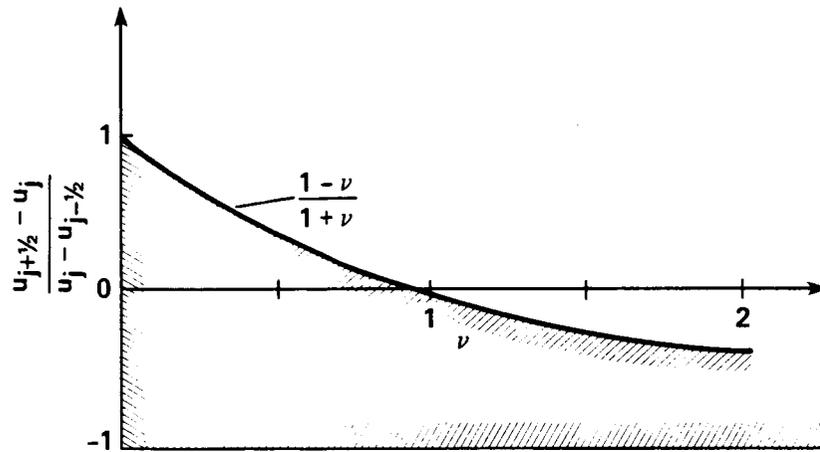


Figure 5. Only schemes in the shaded region satisfy the entropy condition.

Squaring Eq. 4.1 and substituting into Eq. 4.2 gives a discrete entropy statement that may be explicitly calculated. Remarkably, this factors, even without specifying $u_{j+\frac{1}{2}}$.

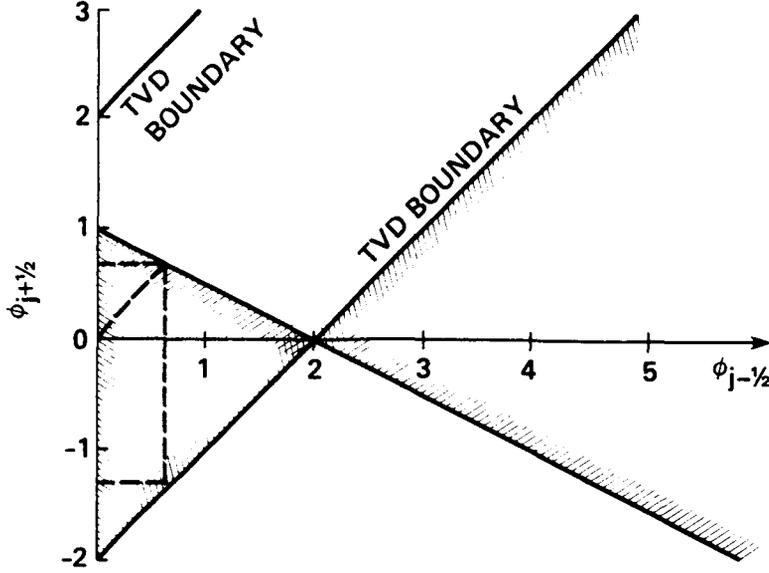


Figure 6. Some TVD schemes satisfy an entropy constraint.

$$(4.3) \quad \nu \left[(u_{j+\frac{1}{2}} - u_j)^2 - (u_j - u_{j-\frac{1}{2}})^2 + \nu (u_{j+\frac{1}{2}} - u_{j-\frac{1}{2}})^2 \right] \leq 0$$

Assume that ν is positive. (The case in which ν is negative is a trivial extension.) It is apparent that the third term will dominate for large enough ν , leading to a kind of CFL restriction on satisfaction of the entropy inequality. The nature of the restriction becomes apparent when Eq. 4.3 is plotted as in Fig. 5. The shaded region indicates satisfaction of the entropy inequality. As in the previous section, the task at hand is to select a maximum value of ν at which to operate. Whatever value is selected, the following inequalities apply.

$$(4.4) \quad -1 \leq \frac{u_{j+\frac{1}{2}} - u_j}{u_j - u_{j-\frac{1}{2}}} \leq \frac{\nu - 1}{\nu + 1}$$

So what does all this have to do with TVD schemes? If the desired difference scheme is given by Eq. 3.5, then appropriate definitions are, as verified by substitution,

$$(4.5) \quad u_{j+\frac{1}{2}} = u_j + \frac{1}{2} \phi_{j+\frac{1}{2}} (u_{j+\frac{1}{2}} - u_j)$$

and an equivalent formula for $u_{j-\frac{1}{2}}$. With these definitions, inequalities (4.4) simplify to

$$(4.6) \quad -1 \leq \frac{\phi_{j+\frac{1}{2}}/r_j}{2 - \phi_{j-\frac{1}{2}}} \leq \frac{\nu - 1}{\nu + 1}$$

These inequalities are plotted in Fig. 6. The smaller the CFL number, the larger the entropy satisfying region. The axes are the same as for Fig. 2, allowing comparison of the TVD region with the entropy satisfying region. In this case entropy satisfying schemes are a subset of TVD schemes.

It is also worth noting that entropy satisfaction alone is insufficient; flux consistency (in the sense of Eq. 2.5) is also required. In Fig. 6 there is an entropy satisfying region to the right of $\phi = 2$ that does not satisfy the flux consistency requirement. Schemes operating in this region are generally unstable.

As explained in the previous section, the case $r_j = 1$ is the most important. This leads to the choice of a rectangular region in Fig. 6 described by

$$(4.7) \quad 0 \leq \phi_{j+\frac{1}{2}} \leq 1 - \nu$$

$$-1 - \nu \leq \frac{\phi_{j+\frac{1}{2}}}{r_j} \leq 1 - \nu$$

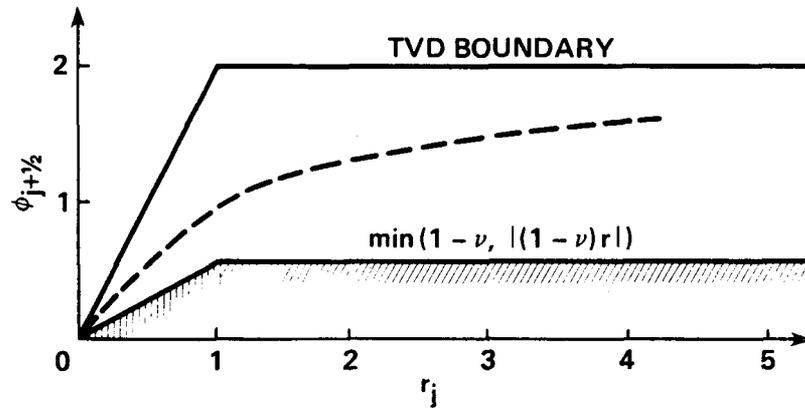


Figure 7. Schemes in the shaded region satisfy an entropy constraint.

This may be compared with Eq. (3.12). Finally these inequalities are plotted in Fig. 7. The shaded region represents a sufficient condition for entropy satisfaction in this case. For comparison, the TVD region is also shown, as is the limiter derived in the previous section. Notice that it lies outside the entropy satisfying region.

On the other hand, it is possible to conclude that the limiter represented by the edge of the entropy satisfying region, will definitely lead to a scheme that satisfies the entropy condition on a cell by cell basis. From the previous section it is easily shown to be a TVD scheme as well. Unfortunately it does not, in general, contain the point $\phi = r_j = 1$ as required for second order accuracy unless $\nu = 0$. It is possible to show (by analysis not

included here) that such a scheme has spatial accuracy of order $\Delta t + \Delta x^2$. This means that it is a first order scheme unless Δt is of the order of Δx^2 . The spatial accuracy, and hence the steady state solution, depends on the time step, a feature that this approach was supposed to avoid. Numerical results are shown in Fig. 8. As expected, the scheme is highly dissipative for a CFL number of $\frac{1}{2}$. It is, however, less dissipative than the first order upwind scheme (and a lot more complicated). That's progress.

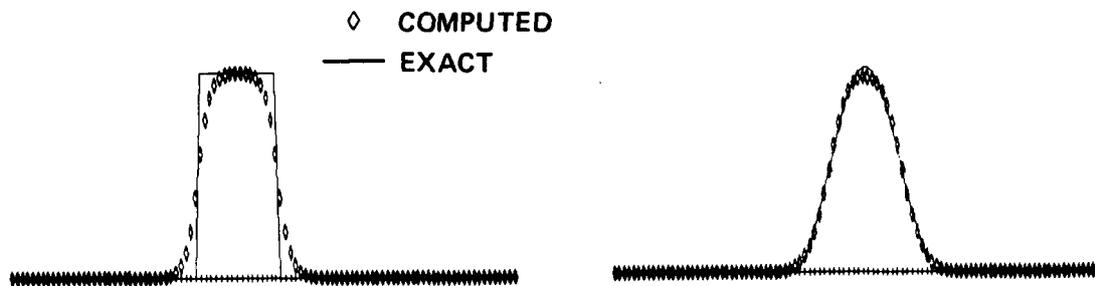


Figure 8. Results for a first order, entropy satisfying scheme .

5. Implicit schemes

Any time a scheme, such as the one just described, shows good features at small CFL numbers that are lost at large ones, it may be possible to avoid the problem using an implicit scheme. In this case an implicit Euler scheme was selected to make use of all the analysis in Sections 3 and 4.

$$(5.1) \quad u^j = u_j - \frac{\nu}{2}(u^{j+1} - u^{j-1})$$

Collecting terms according to time level gives

$$(5.2) \quad u^j + \nu(u^{j+\frac{1}{2}} - u^{j-\frac{1}{2}}) = u_j$$

Recall that the entropy statement must be differenced (integrated) the same way as the PDE. This leads to a slightly different entropy statement:

$$(5.3) \quad (u^j)^2 - u_j^2 + \nu((u^{j+\frac{1}{2}})^2 - (u^{j-\frac{1}{2}})^2) \leq 0$$

Squaring Eq. 5.2 and substituting for u_j^2 in Eq. 5.3 gives, for positive ν ,

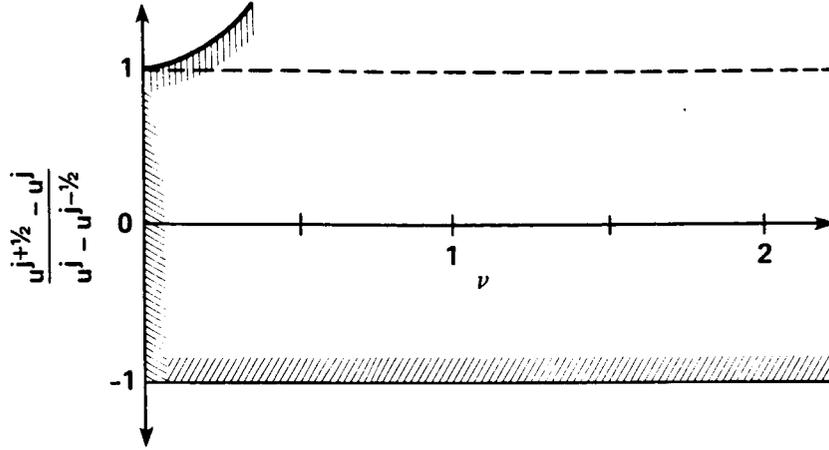


Figure 9. Entropy satisfying regions for implicit scheme.

$$(5.4) \quad (u^{j+\frac{1}{2}} - u^j)^2 - (u^j - u^{j-\frac{1}{2}})^2 - \nu(u^{j+\frac{1}{2}} - u^{j-\frac{1}{2}})^2 \leq 0$$

Notice the similarity to Eq. 4.3. Except for the time level, the only difference is in the sign of the third term. To see what effect this has, inequalities 5.4 are plotted in Fig. 9. The shaded region represents satisfaction of both inequalities. This region is much larger than the equivalent region for explicit Euler time advance (Fig. 5). The entropy inequality will be satisfied, independent of ν , so long as

$$(5.5) \quad -1 \leq \frac{u^{j+\frac{1}{2}} - u^j}{u^j - u^{j-\frac{1}{2}}} \leq 1$$

Following the methodology shown in the previous section leads to the limiter

$$(5.6) \quad \phi^{j+\frac{1}{2}} = \min(1, |r^j|)$$

which is plotted in Fig. 10. Except for the difference in time level, this is the same limiter computed in the previous section if $\nu = 0$. Its advantage occurs for larger values of ν . Notice that central differencing is used except for a small region around the origin. This means that the accuracy of the scheme may (or may not) remain second order accurate at extrema (negative values of r^j). By contrast, the limiter shown in Fig. 3 will always degenerate to first order at extrema.

There is still the problem of computing $u^{j+\frac{1}{2}}$ at the new time step. This can be approximated by a Taylor series type linearization. For example $u^{j+\frac{1}{2}} \approx u_{j+\frac{1}{2}} + \frac{\partial u_{j+\frac{1}{2}}}{\partial u_k} (u^k - u_k)$

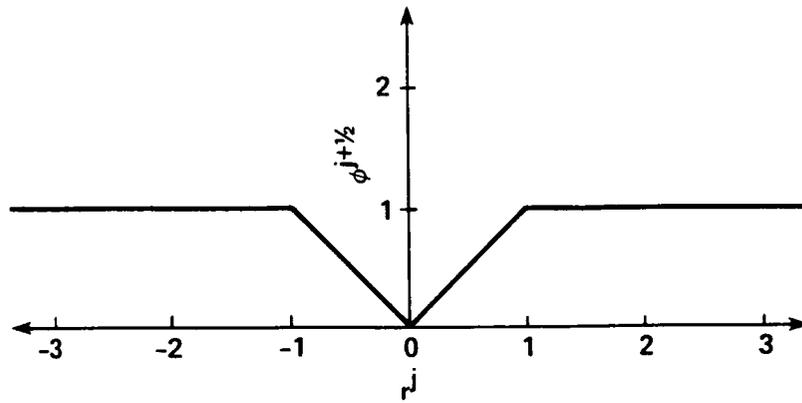


Figure 10. Limiter for implicit TVD scheme.

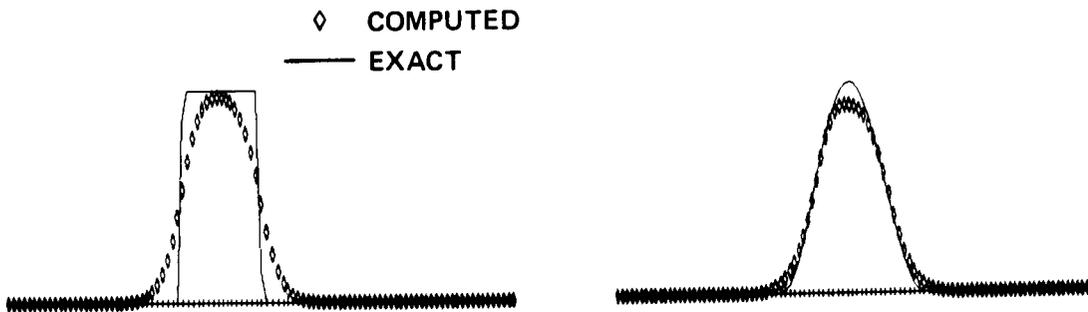


Figure 11. Results for implicit scheme.

where $u_{j+\frac{1}{2}}$ is computed just like $u^{j+\frac{1}{2}}$, but at the old time level. The k subscript is summed, as in tensor notation, over all values of j (most of the terms vanish). A system of linear equations can be formed that has two bands below the diagonal and one above the diagonal. In a more general situation this would be a pentadiagonal. In computing the partials, the only difficulty comes in linearizing the limiter, $\phi(r)$. This must be done piecewise and can be a poor approximation in the transient at high CFL numbers. Numerical results are plotted in Fig. 11. At low CFL numbers the scheme performs very well, producing a TVD result that satisfies the entropy inequality. At high CFL numbers ($\nu \gg 1$) the entropy inequality (and the TVD condition) may not be satisfied in the transient because of errors introduced by the linearization. The scheme remains stable but becomes very diffusive due to the first order accurate time advance. At CFL numbers between $\nu = .5$ and $\nu = 1$ the scheme may actually perform worse than the first order upwind scheme. This is due to the latter schemes special properties for the wave equation at $\nu = 1$.

6. Nonlinear Schemes

It is reasonable to suggest that the wave equation is irrelevant, that it is too far removed from the interesting applications. In spite of this, few schemes have good analysis for anything else. In this case, however, more complete analysis is possible. The case of scalar, nonlinear equations is covered in this section.

Consider the scalar, nonlinear convection equation.

$$(6.1) \quad q_{,t} + f(q)_{,x} = 0$$

For this section, at least, q and f are considered to be scalars. For example, the special case $f(q) = \frac{1}{2}q^2$ is often known as Burger's equation. As in the last section, an implicit difference scheme is selected.

$$(6.2) \quad q^j - q_j + \lambda(f^{j+\frac{1}{2}} - f^{j-\frac{1}{2}}) = 0$$

Instead of ν the definition $\lambda \equiv \frac{\Delta t}{\Delta x}$ is used since the wave speed is no longer constant.

The entropy statement in this case is

$$(6.3) \quad S^j - S_j + \lambda(F^{j+\frac{1}{2}} - F^{j-\frac{1}{2}}) \geq 0$$

Since the generalized entropy S can be a somewhat complicated function of q , it may not be practical to simply substitute Eq. 6.2 into 6.3 as was done for the wave equation. Instead, S_j is expanded in a truncated Taylor series.

$$(6.4) \quad S_j = S^j - S_{,q}^j \Delta q + \frac{1}{2} S_{,qq}(\xi, j) \Delta q^2$$

In Eq. 6.4, Δq is defined by $\Delta q \equiv q^j - q_j$. Also, the third term is evaluated at a time ξ , an unspecified time between the old and new time levels whose existence is guaranteed by the intermediate value theorem. Equation 6.4 is exact, all the uncertainty is tied up in the value of ξ . Substituting this relation into Eq. 6.3 gives

$$(6.5) \quad S_{,q}^j \Delta q + \lambda(F^{j+\frac{1}{2}} - F^{j-\frac{1}{2}}) \geq \frac{1}{2} S_{,qq}(\xi, j) \Delta q^2$$

Since $S(q)$ must satisfy Eq. 2.3 (convexity), the right hand side of Eq. 6.5 is always negative or zero. It is zero at steady state ($\Delta q = 0$). A more restrictive condition (which is easier to implement than Eq. 6.5) is

$$(6.6) \quad S_{,q}^j \Delta q + \lambda(F^{j+\frac{1}{2}} - F^{j-\frac{1}{2}}) \geq 0$$

Notice that Eq. 6.2 gives a relation for Δq . Substituting this into Eq. 6.6 and dividing out λ gives

$$(6.7) \quad -S_{,q}^j (f^{j+\frac{1}{2}} - f^{j-\frac{1}{2}}) + (F^{j+\frac{1}{2}} - F^{j-\frac{1}{2}}) \geq 0$$

Equation 6.7, though correct, is difficult to implement because of terms evaluated between grid points. The solution selected here is to expand such terms in a Taylor series about the point j to two terms. For example the approximation

$$(6.8) \quad f^{j+\frac{1}{2}} \approx f^j + f_{,q}^j \Delta q + \frac{1}{2} f_{,qq}^j \Delta q^2$$

is used. After dropping the obvious terms, what remains is

$$(6.9) \quad (a+b)(F_{,q}^j - S_{,q}^j f_{,q}^j) + \frac{1}{2}(a^2 - b^2)(F_{,qq}^k - S_{,q}^j f_{,qq}^j) \geq 0$$

where $a \equiv q^{j+\frac{1}{2}} - q^j$ and $b \equiv q^j - q^{j-\frac{1}{2}}$. The first term is zero, since $S(q)$ and $F(q)$ must satisfy Eq. 2.4. The second term simplifies substantially using

$$(6.10) \quad F_{,qq} = (F_{,q})_{,q} = (S_{,q} f_{,q})_{,q} = S_{,qq} f_{,q} + S_{,q} f_{,qq}$$

which comes from using Eq. 2.4 and the chain rule. Using Eq. 6.10 to simplify Eq. 6.9 gives the entropy statement,

$$(6.11) \quad \frac{1}{2}(a^2 - b^2) S_{,qq} f_{,q} \geq 0$$

which simplifies further using $S_{,qq} < 0$. The final result is

$$(6.12) \quad \text{sign}(f_{,q})(a^2/b^2 - 1) \leq 0$$

The quantity $f_{,q}$ is just the local wave speed. The term $\text{sign}(f_{,q})$ appears to be required for satisfaction of the discrete entropy inequality. Instead of being postulated for

physical reasons (as in upwind schemes), it appears directly in the algebra. The linearization difficulties posed by such a term are considerable.

In spite of the simplicity of Eq. 6.12, there are a few complications to solving nonlinear problems. First, the discrete entropy inequality is not satisfied exactly, as it was for the wave equation, but only to second order. This is a consequence of approximations such as Eq. 6.8. Second, and more serious, is the problem of sign changes in the local wave speed (shocks and sonic points).

Equation 6.12 must be satisfied at each node point. The value of $q^{j+\frac{1}{2}}$ that is chosen must be compatible with inequalities at point j and at point $j+1$. The important feature to take into account is the signs of the wave speeds at these two points.

The case in which both are positive was covered in the previous section. Equation 6.12 is exact for the wave equation and leads to exactly the same algorithm previously discussed. The case in which both are negative is easily derived by reversing the coordinate system. If, by chance or design, $f_{,q} = 0$ for a particular node, Eq. 6.12 will always be satisfied. The interesting cases are the sign changes, the shock, and sonic point cases.

A sonic point is said to occur if $f_{,q}^j < 0$ and $f_{,q}^{j+1} > 0$. The characteristics are said to be diverging at this point. This is a problem since ϕ_j depends in general on the ratio r_j or r_{j+1} depending on which is the upstream point. At the sonic point neither side can be said to be upstream. In choosing $\phi^{j-\frac{1}{2}}$ it is subtly assumed that $1 \leq \phi^{j+\frac{1}{2}} \leq 2$. In choosing $\phi^{j+\frac{3}{2}}$ it is subtly assumed that $0 \leq \phi^{j+\frac{1}{2}} \leq 1$. The only solution that satisfies entropy inequalities for cells j and $j+1$ is $\phi^{j+\frac{1}{2}} = 1$.

A shock point is said to occur if $f_{,q}^j > 0$ and $f_{,q}^{j+1} < 0$. The characteristics are said to be converging at this point. This case is the most difficult. Since $q(x)$ is discontinuous at the shock, Taylor series approximations such as Eq. 6.8 are not valid. Equation 6.12 must be replaced by two equations, one on each side of the shock

$$(6.13) \quad (F^{j+\frac{1}{2}} - F^j) - S_{,q}^j(f^{j+\frac{1}{2}} - f^j) \geq 0$$

$$(F^{j+1} - F^{j+\frac{1}{2}}) - S_{,q}^{j+1}(f^{j+1} - f^{j+\frac{1}{2}}) \geq 0$$

The endpoint values, $\phi = 0$ and $\phi = 2$, make good guesses because they at least satisfy one of equations 6.13. For the special case of Burger's equation it can be shown that they are the only possible solutions. Furthermore, for some combinations of q^j and q^{j+1} only one of these is valid,

$$(6.14) \quad \phi^{j+\frac{1}{2}} = 1 - \text{sign}(f_{,q}^j + f_{,q}^{j+1})$$

This says that on an interval with a shock in it, $u^{j+\frac{1}{2}} = u^j$ or $u^{j+\frac{1}{2}} = u^{j+1}$. The choice is determined by the sign of the shock speed, which is consistent with characteristic theory. Assuming that Eq. 6.14 is valid in general, the limiter formula's for a scalar nonlinear equation are summarized in Table 1.

Table 1 - The Function $\phi(r)$

$\phi^{j+\frac{1}{2}} =$	$f_{,q}^j \leq 0$	$f_{,q}^j > 0$
$f_{,q}^{j+1} \leq 0$	$2 - \min(1, r^{j+1} ^{-1})$	$1 - \text{sign}(f_{,q}^j + f_{,q}^{j+1})$
$f_{,q}^{j+1} > 0$	1	$\min(1, r^j)$

For the special case of Burger's equation, a suitable choice of entropy and entropy flux is $S(q) = q^2$, $F(q) = \frac{2}{3}q^3$. An interesting initial condition is $q(x) = \sin(x)$ on the interval $x = [0, 2\pi]$. This case was computed for several values of λ . Figure 12 shows the results for an initial CFL value of $\frac{1}{10}$. The computed values are shown as symbols whereas the exact values are shown as a solid line. Also shown is $\phi(x)$ at the beginning and end of the calculation. Notice that central differencing is used in expansion regions and second order upwind differencing is used in compression regions. At the shock a first order upwind scheme is employed. Once the shock has formed, all of the dissipation occurs at the shock, a feature of the PDE.

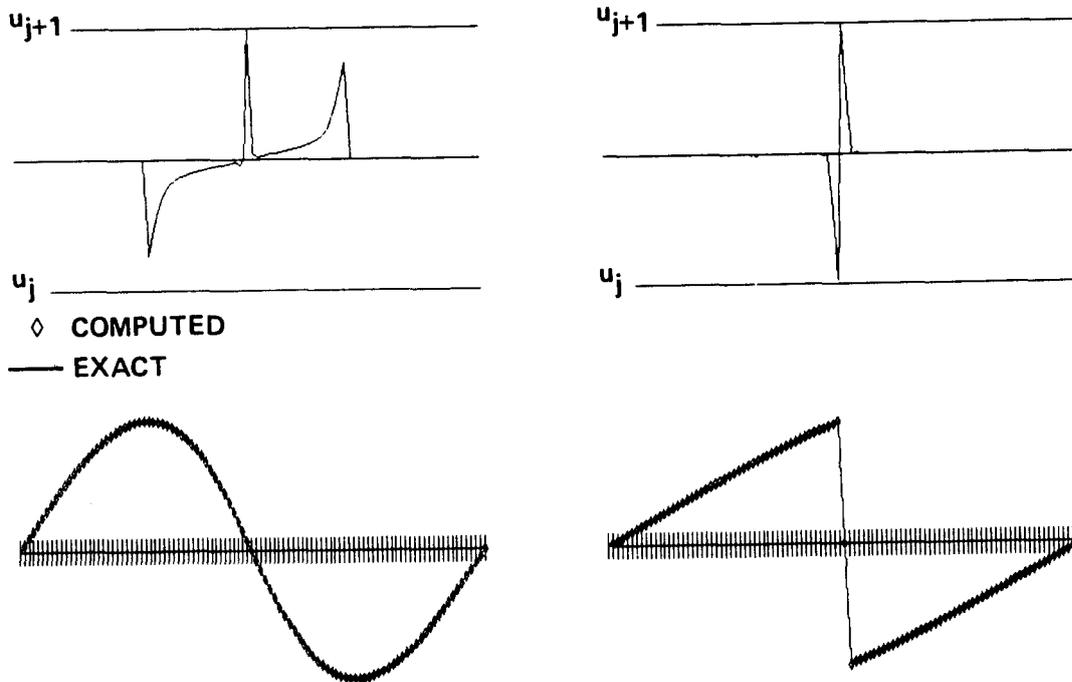


Figure 12. The differencing is a function of the solution.

There are a few anomalies, not apparent from Fig. 12. The limiters are computed in such a way that only a single interval is allowed for the shock. Yet, in this worst case

scenario, the shock lands on a grid point. What appears to be a two interval shock is treated by the scheme as a shock and a strong compression. If, due to roundoff, the shock is initially moving to the right, the center point will rise above zero. This will result in the shock moving to the right on the next step. After the solution matures, there is no visible difference between the shock and the compression. Another problem is that pure central differencing is neutrally stable. Occasionally, rounding errors will result in a visible excursion of the limiter away from central differencing even in the smooth regions. These excursions are very small, however, and the solution is not noticeably affected.

At moderate CFL numbers ($\lambda \approx 1$) the scheme suffers from minor time accuracy problems in the transient. This is a result of the first order time advance. At high CFL numbers ($\lambda \gg 1$), the scheme loses the TVD property in the transient because of linearization errors. It remains stable, however, and achieves the correct steady state solution at CFL numbers of 1000.

7. Conclusions

Initially this work was undertaken to explain artificial dissipation ("smoothing") in physical terms. The explanation is that smoothing is a numerical artifice to satisfy the discrete entropy inequality in a global sense. This ensures stability (given suitable boundary conditions) since the total entropy within the domain must be non-diminishing. Entropy in a region reaches a maximum when the solution is a constant over the entire domain. Normally a balance is reached before this happens, but in any case the solution cannot blow up. If a global entropy inequality is satisfied while a local one is not, there will be several cells (for example, near shocks) that destroy entropy. This allows the oscillations so familiar with linear smoothing. Physically, the solution is only meaningful if integrated over a collection of cells that, together, satisfy the entropy inequality. Near shocks, this may require several cells to be lumped together and averaged. Thus the effective grid near a shock may be coarser than it appears if the smoothing does not satisfy a cell entropy inequality.

It should also be mentioned that smoothness of a solution is no guarantee that a local entropy inequality is satisfied. The author has observed smooth solutions of the quasi-one-dimensional Euler equations for which every cell in the supersonic region consumes entropy. Such solutions do not correspond to flows of very viscous fluid as is often thought. That they do not correspond to any physical fluid is evidenced by the fact that the second law of thermodynamics is violated over much of the domain.

The flux consistency condition puts bounds on what happens between grid points. In this work a piecewise constant assumption is used to compute the discrete entropy inequality. The evaluation of flux between gridpoints must be consistent with this assumption. The flux consistency condition used here is not unique, just sufficient. It is only valid if the piecewise constant assumption is used.

The purpose of this work is not so much to recommend a particular scheme as it is to recommend a method of analysis based on satisfying the entropy inequality and a certain flux consistency condition. The scheme derived here is offered as proof that schemes which satisfy these two conditions do exist.

This work has shown that the TVD property is insufficient to guarantee a physical solution, either in the steady state or in the transient. To provide such a guarantee, a scheme must satisfy the discrete entropy inequality for time dependent problems, or the semi-discrete entropy inequality for steady state problems. Schemes that satisfy a discrete entropy inequality may need to be implicit to achieve second order accuracy in space.

References

1. **P.K. Sweby**, *High Resolution Schemes Using Flux Limiters for Hyperbolic Conservation Laws*, SIAM Journal Of Numerical Analysis, 5, (1984), pp. 995-1011.
2. **P.L. Roe**, *Numerical Algorithms for the Linear Wave Equation*, Royal Aircraft Establishment Technical Report 81047, 1981.
3. **B. van Leer**, *Towards the Ultimate Conservative Difference Scheme, II. Monotonicity and Conservation Combined in a Second Order Scheme*, Journal of Computational Physics, 14 (1974), pp. 361-370.
4. **S. Chakravarthy and S. Osher**, *High Resolution Applications of the Osher Upwind Scheme for the Euler Equations*, AIAA paper 83-1943 presented at 6th CFD conference 1983.
5. **P.D. Lax and B. Wendroff**, *Systems of Conservation Laws*, Communications on Pure and Applied Mathematics, 13 (1960), pp. 217-237.
6. **J. B. Goodman and R. J. LeVeque**, *On the Accuracy of Stable Schemes for 2D Scalar Conservation Laws*, Mathematics of Computation, 171 (1985), pp 15-21.
7. **P.D. Lax**, *Hyberbolic Systems Of Conservation Laws and the Mathematical Theory of Shock Waves*, CBMS Regional Conference Series in Applied Mathematics, Philadelphia, 1973.
8. **A. Harten**, *High Resolution Schemes for Conservation Laws*, Journal of Computational Physics, 49 (1983), pp 357-393.

1. Report No. NASA TM-88371		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle SMOOTHING AND THE SECOND LAW				5. Report Date September 1986	
				6. Performing Organization Code	
7. Author(s) Marshal L. Merriam				8. Performing Organization Report No. A-86424	
9. Performing Organization Name and Address Ames Research Center Moffett Field, CA 94035				10. Work Unit No.	
				11. Contract or Grant No.	
				13. Type of Report and Period Covered Technical Memorandum	
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546				14. Sponsoring Agency Code 505-60	
				13. Type of Report and Period Covered	
15. Supplementary Notes Point of Contact: Marshal L. Merriam, Ames Research Center, M/S 202A-1 Moffett Field, CA 94035 (415) 694-6417 or FTS 464-6417					
16. Abstract <p>The technique of obtaining second order, oscillation free, total variation diminishing (TVD), scalar difference schemes by adding a limited diffusive flux ("smoothing") to a second order centered scheme is explored. It is shown that such schemes do not always converge to the correct physical answer. The approach presented here is to construct schemes that numerically satisfy the second law of thermodynamics on a cell by cell basis. Such schemes can only converge to the correct physical solution and in some cases can be shown to be TVD. An explicit scheme with this property and second order spatial accuracy was found to have an extremely restrictive time step limitation ($\Delta t < \Delta x^2$). Switching to an implicit scheme removed the time step limitation.</p>					
17. Key Words (Suggested by Author(s)) Total variation diminishing (TVD) Dissipation Entropy			18. Distribution Statement Unlimited Subject Category - 64		
19. Security Classif. (of this report) Unclassified		20. Security Classif. (of this page) Unclassified		21. No. of Pages 23	22. Price* A02